

Open Research Online

The Open University's repository of research publications
and other research outputs

Visual Analytics for Health Monitoring and Risk Management in CARRE

Journal Item

How to cite:

Zhao, Youbing; Parvinzamid, Farzad; Wei, Hui; Liu, Enjie; Deng, Zhikun; Dong, Feng; Third, Allan; Lukoševičius, Arnas; Marozas, Vaidotas; Kaldoudi, Eleni and Clapworthy, Gordon (2016). Visual Analytics for Health Monitoring and Risk Management in CARRE. E-Learning and Games; 10th International Conference, Edutainment 2016, Hangzhou, China, April 14-16, 2016, Revised Selected Papers, 9654 pp. 380–391.

For guidance on citations see [FAQs](#).

© 2016 Springer International Publishing Switzerland



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

http://dx.doi.org/doi:10.1007/978-3-319-40259-8_33

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Visual Analytics for Health Monitoring and Risk Management in CARRE

Youbing Zhao¹(✉), Farzad Parvinzmir¹, Hui Wei¹, Enjie Liu¹, Zhikun Deng¹,
Feng Dong¹, Allan Third², Arūnas Lukoševičius³, Vaidotas Marozas³,
Eleni Kaldoudi⁴, and Gordon Clapworthy¹

¹ University of Bedfordshire, Luton LU1 3JU, UK
{youbing.zhao,farzad.parvinzmir,hui.wei,enjie.liu,zhikun.deng,feng.dong,
gordon.clapworthy}@beds.ac.uk

² Knowledge Media Institute, The Open University, Milton Keynes, UK
allan.third@open.ac.uk

³ Biomedical Engineering Institute, Kaunas University of Technology,
Kaunas, Lithuania
{arunas.lukosevicius,vaimaro}@ktu.lt

⁴ Physics of Medical Imaging and Telemedicine, School of Medicine,
Democritus University of Thrace Dragana, Alexandroupoli, Greece
kaldoudi@med.duth.gr

Abstract. With the rise of wearable sensor technologies, an increasing number of wearable health and medical sensors are available on the market, which enables not only people but also doctors to utilise them to monitor people's health in such a consistent way that the sensors may become people's lifetime companion. The consistent measurements from a variety of wearable sensors implies that a huge amount of data needs to be processed, which cannot be achieved by traditional processing methods. Visual analytics is designed to promote knowledge discovery and utilisation of big data via mature visual paradigms with well-designed user interactions and has become indispensable in big data analysis. In this paper we introduce the role of visual analytics for health monitoring and risk management in the European Commission funded project CARRE which aims to provide innovative means for the management of cardiorenal diseases with the assistance of wearable sensors. The visual analytics components of timeline and parallel coordinates for health monitoring and of node-link diagrams, chord diagrams and sankey diagrams for risk analysis are presented to achieve ubiquitous and lifelong health and risk monitoring to promote people's health.

Keywords: CARRE · Visual analytics · Wearable sensor · Health monitoring · Risk management

1 Introduction

The widespread use of wearable monitoring devices and mobile apps makes ubiquitous capture of life-logging personal health data a reality. Effective collection

of long-term health-status data is valuable for clinical decisions and leads to strengthened interdisciplinary healthcare research and collaboration in supporting innovative medical care. With different types of wearable sensor the healthcare platform can aggregate heterogeneous health and medical data for health monitoring, disease prediction and risk management in a ubiquitous, personalised and continuous manner.

In this paper we present the ongoing work of visual analytics in the CARRE project [1] – Personalized Patient Empowerment and Shared Decision Support for Cardiorenal Disease and Comorbidities – funded by the 7th Framework Programme of the European Commission, which aims to provide innovative means for the management of cardiorenal diseases with the assistance of wearable sensors.

Wearable device based health data collection tools traditionally refers to use of medical devices to monitor medical data, such as heart rate, blood pressure, glucose, etc. Recently, the use of wearable devices in life-logging data collection mainly indicates the record of some personal physical activity data. CARRE uses Fitbit [2], Withings [6] and iHealth citeiHealth devices to record steps, distance, calories, heart rate, sleep quality, blood pressure, weight, etc.

The target of CARRE is to provide personalised empowerment and shared decision support for cardiorenal disease, which is the condition characterised by simultaneous kidney and heart disease while the primarily failing organ may be either the heart or the kidney. In CARRE, sources of medical and other knowledge will be semantically linked with sensor outputs to provide clinical information personalised to the individual patient, so as to be able to track the progression and interactions of comorbid conditions. The ultimate goal is to provide the means for patients with comorbidities to take an active role in care processes, including self-care and shared decision-making, and also to support medical professionals in understanding and treating comorbidities via an integrative approach.

The CARRE repository acts as the central point of information storage for all CARRE applications. It conforms to the principles of the Semantic Web and the guidelines of Linked Data. The Linked Data guidelines can be summarized as follows:

Information stored in the CARRE repository consists of ‘RDF triples’. RDF is a standard format for representing semantic data on the Web; an item of RDF data is a triple, which corresponds to a statement of the form ‘subject predicate object’. Each term is a URI, often drawn from a standard vocabulary or ontology, making it easy to link triples from different sources – to allow Linked Data. RDF can be accessed through a SPARQL endpoint: SPARQL is a query language, much like SQL in syntax. The triples stored in the CARRE repository are either public or private. For private data, data privacy and security mechanisms have been deployed.

The CARRE repository [21] stores general medical knowledge relating to risk associations, evidence and observables, and is available for public querying without authentication, as it contains no personally identifying data for any

patient and serves as a general-purpose resource for medical knowledge in a semantic format.

In CARRE the risk model is a large semantic graph structure data consisting of interlinked entities, such as risk elements and risk evidence, that are either related to ground knowledge in cardio-renal disease and comorbidities (symptoms, diseases, risk factors, treatments, medical evidence source data, educational content, etc.) or personalised to each patient (patient demographics, medical history, sensor data, lifestyle data, etc.) [13]. The data structure of the risk factor repository is shown in Fig. 1.

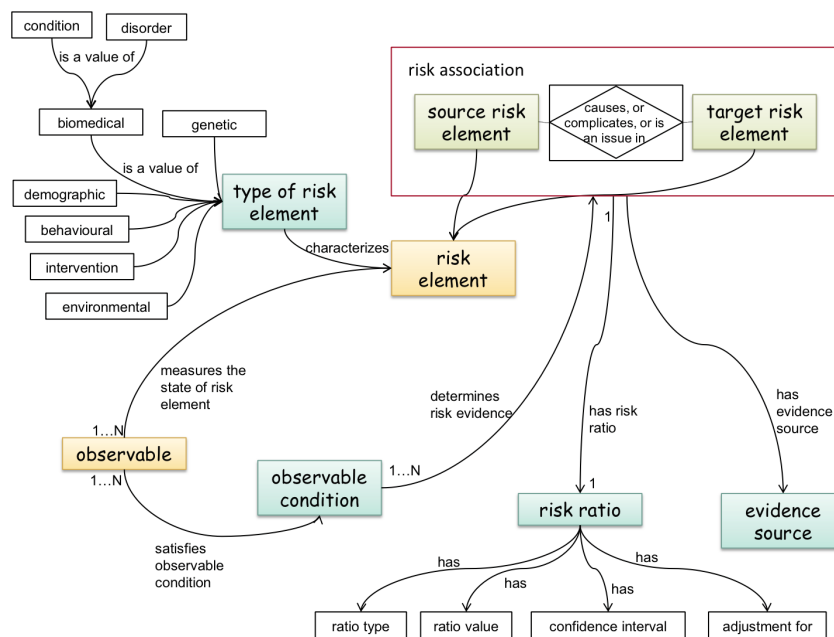


Fig. 1. The risk model structure in the CARRE risk repository

The key concepts in the CARRE risk factor network are defined as follows:

- **Risk Element:** Risk elements include all the disorders/diseases involved in the comorbidity under discussion as well as any other risk causing agent, e.g. demographic (e.g. age, sex, race), genetic (gene polymorphisms), behavioural (e.g. smoking, physical exercise), environmental (e.g. air pollution, allergens) or even an intervention (e.g. pharmaceutical substances, contrast agents).
- **Risk Association:** The association of one risk element as the risk source with another risk element, which is the negative outcome under certain conditions, is a 'risk association'. A source risk element can be associated to a target risk element with more than one risk association.

- **Risk Ratio:** The association is always accompanied by the likelihood of the negative outcome occurring. This likelihood is expressed as a ‘risk ratio’, which is the ratio of the probability of the negative outcome when the person is exposed to the risk agent over the probability of the negative outcome when the person is not exposed to the risk agent.
- **Risk Observable:** In a risk association the prerequisite circumstances relate directly to the existence of the risk agent (source risk target) and/or its severity, and/or any other specific conditions. These are reported via certain ‘observables’, that is, physical variables that can be measured or otherwise ascertained (e.g. biomarkers, biometric variables, biological signals and other non-biological factors, e.g. environmental).

Currently there are 98 risk factors, 53 risk elements, 253 risk evidences, 63 observables and 60 evidence sources in the CARRE risk data repository.

The risk model provided by CARRE can also be applied to other platforms for effective health monitoring and disease diagnosis. For example, the bridging of the CARRE risk model and the life-logging platform MyHealthAvatar [4,5] is under investigation.

To view and understand both the health status data and the risk data, visual analytics is indispensable in CARRE to provide patients and clinicians the ability to visualise, understand and interact with this linked knowledge and also take advantage of personalised empowerment services. The aim is to help medical professionals to better understand the individual patient’s disease development and help patients to understand their own disease development, which in turn assists them to adhere to the self-management plan.

This paper presents the first stage of the CARRE Visual Analytics implementation: timeline and parallel coordinates are employed for health monitoring and correlation analysis, node-link diagrams, chord diagrams etc. are chosen for risk factor visualisation and risk analysis. A preliminary version of personalised risk visualisation and disease propagation simulation is demonstrated.

The paper is organised as follows: Sect. 2 introduces related work in health data visual analytics; Sect. 3 discusses major visual analytics components of CARRE for health monitoring and risk management and Sect. 4 concludes with the summary and future work.

2 Related Work

Healthcare has been an important research and application field of data analysis and visualisation for several decades [8]. Behind healthcare visual analytics, much of the focus is on the visualisation of electronic health records (EHRs). [7] gives a detailed review of the related work, categorising by individual patients or group of patients. In each category, the work is further divided by visual analytics of time series data or status at a certain time point. [23] also presents a systematic review of innovative visual analytics approaches that have been proposed to illustrate EHR data.

Lifelines [18] is a pioneer work in visualisation of individual patient records, which provides a general visualisation environment for problems, diagnoses, test results or medications using timelines.

Lifelines2 [22] provides visualisation of temporal categorical data across multiple records, which is better for a doctor to view to discover and explore patterns across these records to support hypothesis generation, and find cause-and-effect relationships in a population.

LifeFlow [25] and EventFlow [16] are tools for event sequence analytics for a group of patients. They extract and highlight the common event sequence from patient records.

Outflow [24] and DecisionFlow [10] uses Sankey diagrams [19] to help to visualise and analyse the causal relationships of events in complex event sequences.

The existing work has made a detailed visual analytics research mostly on EHRs and event analysis. While the work on predictive visual analytics of healthcare data is highly valuable [17], it is still rare due to its complexity.

Personal health information has been increasingly collectible and accessible in the information era. With the trend of “predictive, preemptive, personalized and participative” healthcare [20], more personalised data is desired for predictive analysis of medical care. In addition, with the development of wearable sensor technologies, there is a vision of the merging of health sensors and medical sensors, which will bring the fusion of personal health information and the medical sensor data in the future. Together with the clinical information that has long played the major role in health and medical decision making, personal health information can introduce more added value for health monitoring and medical decision making.

CARRE can not only directly access personal health and lifestyle data from devices such as Fbit, Withings and iHealth, but also access data from multiple heterogeneous data sources via Microsoft HealthVault [3]. Together with proven risk models [13] extracted from medical literatures, CARRE aims to provide personalised risk management and analysis, which is a key difference from the existing healthcare visual analytics systems.

3 CARRE Visual Analytics

Visual analytics combines automated analysis techniques with interactive visualisations for an effective understanding, reasoning and decision-making on the basis of very large and complex datasets [14]. It is designed to promote knowledge discovery and utilisation of big data via effective visual paradigms and well-designed user interactions. In visual analytics, visualisation becomes the medium of an interactive analytical process, where humans and computers cooperate using their respective distinct capabilities for data processing and visual recognition for the most effective results. Thus, in visual analytics, user interaction constitutes the key for a user-centred data exploration. Visual analytics is an indispensable technology for information processing in the big data era.

In CARRE the data can be generally categorised into fitness data collected from sensors/apps, biomarker data from personal electronic health records

(PHR) and risk factor data extracted from medical literatures. The role of visual analytics is to visualise health data, risk factor data and the integrated visual analysis of health data and risk factor data. In the current first stage implementation, CARRE provides web-based components for interactive health data visualisation and risk analysis, including healthline and parallel coordinates for fitness and biomarker data, node-link diagram, chord diagram and sankey diagram for risk factor data visual analysis and a preliminary experiment on personalised risk visualisation and disease progression simulation.

3.1 Fitness and Biomarker Data Visualisation

Healthline. Lifestyle, health, fitness and medical data are inherently time dependent. To visualise time-varying data, a linear form timeline is a natural choice. A healthline is a special form of timeline to visualise multiple variables of continuous fitness statistics and biomedical markers which may cover a long period. Data trends can be observed from the variable curves and data correlations may be discovered by comparison of the data curves of the multi-variables. As the data records may cover a long time range, interactive techniques such as zooming and overview+details [9] are employed in the healthline visualisation. The users can also select the variables they are interested in the variable legend list to perform a user defined variable filtering. Figure 2 shows multiple biomarkers visualised in the interactive healthline in CARRE.

Parallel Coordinates. The technique of parallel coordinates is an approach for visualising multiple quantitative variables using multiple axes which are placed parallel to each other in the most common case [12]. The advantage of parallel coordinates is that it supports visualisation of multiple variables and correlation



Fig. 2. The healthline visualises personal fitness and biomarker data

between attributes can be discovered by certain visualisation patterns. It is a common technique of visualising high-dimensional data and analyzing multivariate data. In CARRE there are a number of fitness and biomarker variables and both patients and medical practitioners like to view or study correlations among different variables. Consequently, parallel coordinates is chosen for multi-variable correlation visualisation analysis of fitness and biomarker data. An example view of the parallel coordinate view is shown in Fig. 3 where negative correlations can be found between walking minutes and blood pressure as well as BMI (Body Mass Index).

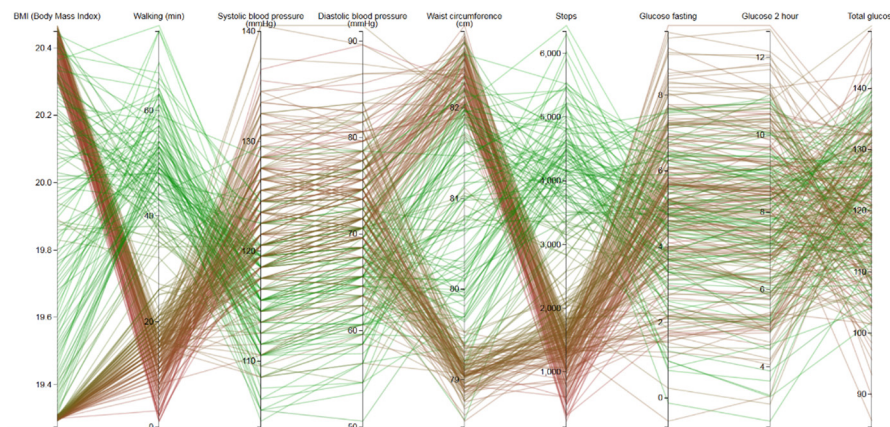


Fig. 3. Fitness and biomarker visualisation and correlation analysis based on Parallel Coordinates

3.2 Risk Factor Visualisation

As introduced in Sect. 1, the risk factor data in the CARRE repository is essentially a graph whose nodes are risk elements with multiple attributes attached, such as risk observables. Each directed edge represents a risk factor directed from the source to the target risk element. In CARRE node-link diagrams as well as other graph visualisation techniques, such as chord diagrams [11], sankey diagrams [19], etc., are used to visualise the risk factor graph. In this paper we focus on the node-link diagram and chord diagram which can be used by professionals and patients to view the risks.

Node-Link Diagram. Node-link diagram [15] is one traditional technique to visualise graph data structure visually. Figure 4 is a force-directed node-link diagram visualisation of all 53 risk elements and 98 risk factors in the current CARRE repository. The node-link diagram clearly visualises risk associations and promotes studying and understanding of the risk factor data base. Through

the visual analytic interface, users are able to explore different diseases by dragging them to the centre to clearly see the relationships. By interactively viewing the graph, a patient also understands their risks in a more intuitive manner.

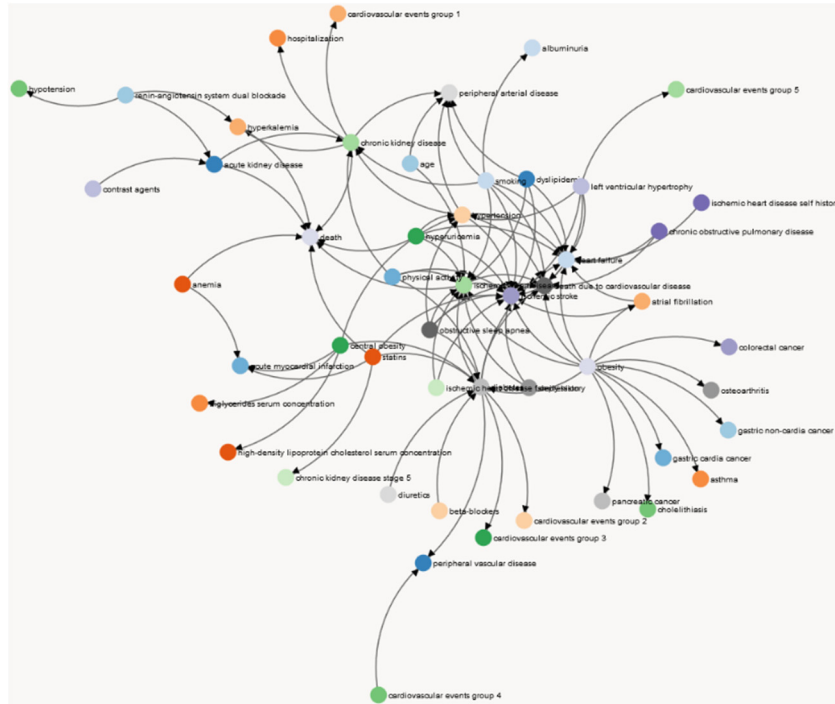


Fig. 4. Force-directed layout based node-link diagram of risk elements and risk factors in CARRE: the whole risk factor database

Chord Diagram. The disadvantage of the node-link diagram is that without proper handling, when the number of nodes and links increase, the visualisation will become increasingly messy for effective recognition by human beings, as obvious in Fig. 4. Though filtering may be applied based on the conditions of a particular patient, it does not help for the visual analysis of the whole risk factor database. Fortunately there are some network visualisation techniques to alleviate the problem, such as the chord diagram [11], as shown in Fig. 5.

The benefits of the chord diagram are that all the nodes are arranged on a circle and the edges from one node are grouped and bundled, which reduces the hairball problems which occur in the node-link diagram. With proper mouse hover interactions all the edges from or to one node can be highlighted, thus making the observation of the connections from or to one node much easier.

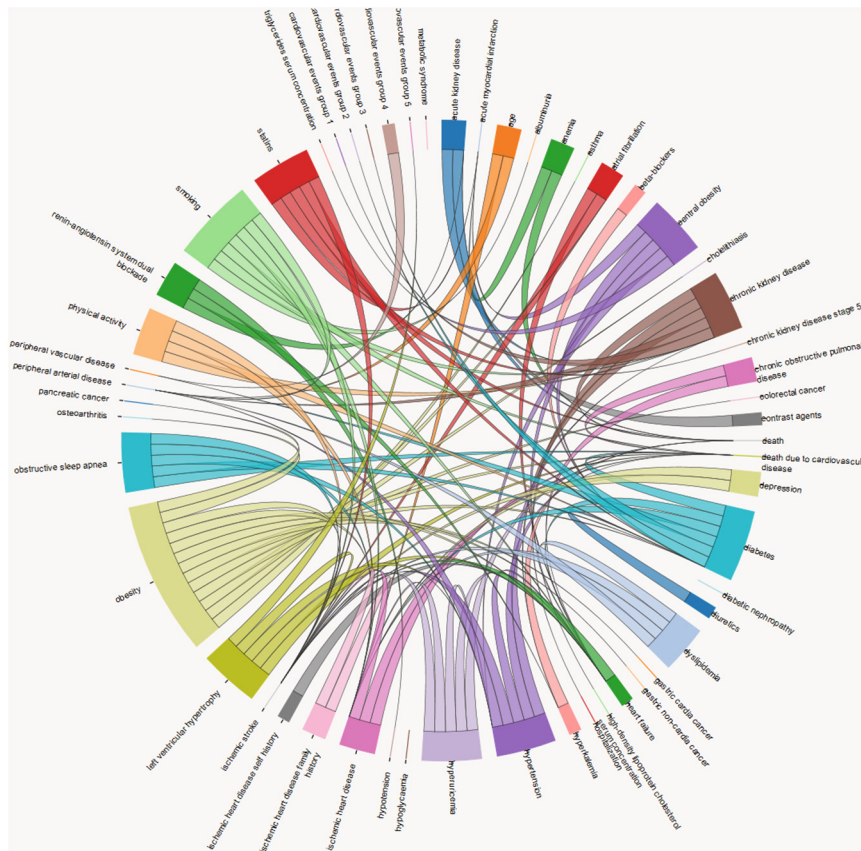


Fig. 5. Chord diagram of risk elements and risk factors in CARRE

The chord diagram clearly visualises the relationships of all risk elements in the repository and is particularly useful when professionals check and insert new risks.

3.3 Visual Analytics of Personalised Risks and Disease Progressions

The ultimate goal of CARRE is to integrate the sensor data and the risk factor database to promote risk discovery, prediction and management, and individual health monitoring. In the first stage of CARRE visual analytics, we also made a prototype of interactive risk visualisation based on individual conditions and inputs. As shown in Fig. 6, the node-link diagram only shows the related risk elements and risk factors based on the patient’s profile, thus greatly reducing the complexity the original diagram. Moreover, the risk associations update dynamically based on changes in the patient’s conditions. Currently, this is achieved by the interactive adjustment of some of the fitness and biomarker data. For

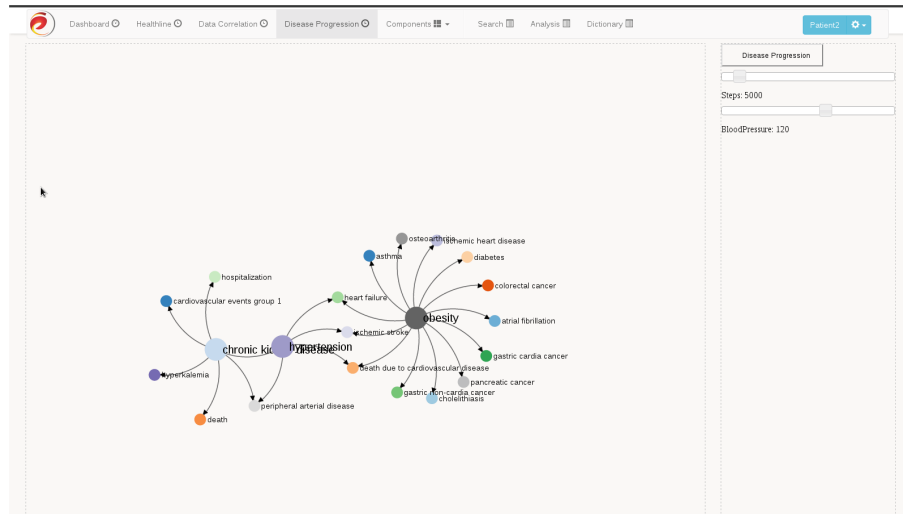


Fig. 6. Interactive risk visual analysis: risks filtered and changed according to the conditions of a particular patient

example, if the blood pressure drops to the normal range, the hypertension risk element will disappear. In another example if the user walks more, the obesity risk element and all risk factors related to obesity will disappear.

4 Conclusions and Future Work

With the increasing popularity of wearable sensors, people are more interested in studying the applications of the sensors for lifestyle tracking and decision support for professional clinicians. While the amount of data collected by wearable sensors is huge, without visual analytics it is almost impossible to carry out any effective data analysis. This paper introduces in particular the role of visual analytics for exploiting sensor data for health monitoring and risk management in CARRE. Multiple variable time-dependant data visualised in a linear healthline helps to study and analyse fitness and biomarker data in CARRE, especially when proper user interaction techniques are incorporated. Parallel coordinates are very useful for correlation analysis of the fitness and biomarker samples. The network of risk elements and risk factors can be visualised with node-link diagrams and chord diagrams, with the latter preferable when the number of nodes becomes large. Moreover, a prototype of individual patient based interactive risk visualisation, prediction and management is presented to show the future directions of our work. In conclusion, interactive visual analytics is critical and effective in sensor-assisted health risk management and analysis; the future work will focus on integrating the risk model more closely with the process of the individualised and dynamic risk management, prediction and visualisation.

Acknowledgments. CARRE project is funded by the European Commission's 7th Framework Programme – ICT under agreement FP7-ICT-611140. The related MyHealthAvatar project is funded by the European Commission's 7th Framework Programme – ICT under agreement FP7-ICT-2011-9. We would like to thank the European Commission for the funding and thank the project officers and reviewers for their indispensable support for both of the projects.

References

1. CARRE Project. <http://www.carre-project.eu/>. Accessed 06 Jan 2016
2. Fitbit. <http://www.fitbit.com/>. Accessed 06 Jan 2016
3. HealthVault. <https://www.healthvault.com/>. Accessed 06 Jan 2016
4. Platform, M.: <http://www.myhealthavatar.org/>. Accessed 06 Jan 2016
5. MyHealthAvatar Project: <http://www.myhealthavatar.eu/>. Accessed 06 Jan 2016
6. Withings. <http://www.withings.com/>. Accessed 06 Jan 2016
7. Rind, A., Wang, T.D., Aigner, W., Miksch, S., Wongsuphasawat, K., Plaisant, C., Shneiderman, B.: Interactive information visualization to explore and query electronic health records. *Found. Trends HCI* **5**(3), 207–298 (2011)
8. Reddy, C.K., Aggarwal, C.C.: *Healthcare Data Analytics*. Chapman & Hall, CRC, Boca Raton (2015)
9. Cockburn, A., Karlson, A., Bederson, B.B.: A review of overview+detail zooming and focus+context interfaces. *ACM Comput. Surv. (CSUR) Surv.* **41**(1), 1–31 (2009)
10. Gotz, D., Stavropoulos, H.: Decisionflow: visual analytics for high-dimensional temporal event sequence data. *IEEE Trans. Vis. Comput. Graph.* **20**(12), 1783–1792 (2014)
11. Holten, D.: Hierarchical edge bundles: visualization of adjacency relations in hierarchical data. *IEEE Trans. Vis. Comput.* **12**(5), 741–748 (2006)
12. Inselberg, A., Dimsdale, B.: Parallel coordinates: a tool for visualizing multi-dimensional geometry. In: *Proceedings of the 1st IEEE Symposium on Visualization*, pp. 361–378. IEEE (1990)
13. Kaldoudi, E., Third, A., Gotsis, G., Roumeliotis, S., Karvelyte, N., Rimseviciu, L., Pafili, K., Papazoglou, D., Juozalenaite, G., Semertzidou, E., Visockiene, Z., Zigerido, K.: CARRE Deliverable D.2.2: Functional Requirements and CARRE Information Model. <https://www.carre-project.eu/project-info/deliverables/> (July 2015). CARRE: Personalized Patient Empowerment and Shared Decision Support for Cardioresenal Disease and Comorbidities, FP7-ICT-611140
14. Keim, D., Kohlhammer, J., Ellis, G., et al.: *Mastering the Information Age - Solving Problems with Visual Analytics*. Eurographics Association, Goslar (2010)
15. Liu, S., Cui, W., Wu, Y., Liu, M.: A survey on information visualization: recent advances and challenges. *Vis. Comput.* **30**(12), 1373–1393 (2014)
16. Monroe, M., Lan, R., Lee, H., Plaisant, C., Shneiderman, B.: Temporal event sequence simplification. *IEEE Trans. Vis. Comput. Graph.* **19**(12), 2227–2236 (2013)
17. Groves, P., Basel Kayylai, D.: *The big-data revolution in US health care: Accelerating value and innovation* (2013)
18. Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., Shneiderman, B., Colorado, K.P.: Lifelines: using visualization to enhance navigation and analysis of patient records. In: *Proceedings of the 1998 American Medical Informatic Association Annual Fall Symposium*, pp. 76–80 (1998)

19. Riehmann, P., Hanfler, M., Froehlich, B.: Interactive sankey diagrams. In: Proceedings of IEEE Symposium on Information Visualization InfoVis 2005, pp. 233–240. IEEE (2005)
20. Shneiderman, B., Plaisant, C., Hesse, B.W.: Improving health and healthcare with interactive visualization methods. *IEEE Comput.* **46**(1), 58–66 (2013)
21. Third, A., Kaldoudi, E., Gotsis, G., Roumeliotis, S., Pafili, K., Domingue, J.: Capturing scientific knowledge on medical risk factors. In: K-CAP2015: 8th International Conference on Knowledge Capture. ACM (2015)
22. Wang, T.D., Plaisant, C., Shneiderman, B., Spring, N., Roseman, D., Marchand, G., Mukherjee, V., Smith, M.: Temporal summaries: Supporting temporal categorical searching, aggregation and comparison. *IEEE Trans. Vis. Comput. Graph.* **15**(6), 1049–1056 (2009)
23. West, V.L., Borland, D., Hammond, W.: Innovative information visualization of electronic health record data: a systematic review. *J. Am. Med. Inform. Assoc.* **22**(2), 330–339 (2015)
24. Wongsuphasawat, K., Gotz, D.: Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization. *IEEE Trans. Vis. Comput. Graph.* **18**(12), 2659–2668 (2012). doi:[10.1109/TVCG.2012.225](https://doi.org/10.1109/TVCG.2012.225)
25. Wongsuphasawat, K., Guerra Gómez, J.A., Plaisant, C., Wang, T.D., Taieb-Maimon, M., Shneiderman, B.: Lifeflow: visualizing an overview of event sequences. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI 2011, NY, USA, pp. 1747–1756. ACM, New York (2011)